Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

THE MATHEMATICS TEACHER

EDITED BY
W. H. METZLER

ASSOCIATED WITH
EUGENE R. SMITH HARRY D. GAYLORD
GEO. GAILEY CHAMBERS WILLIAM E. BRECKENRIDGE

VOLUME XI JUNE, 1919 NUMBER 4

TESTS OF MATHEMATICAL ABILITY—THEIR SCOPE AND SIGNIFICANCE.

By Agnes L. Rogers.

Measurement in the field of algebra and geometry is of such recent origin, it may at first glance seem somewhat premature to survey the work so far accomplished. It was only in 1914 that the first tests of ability in high school mathematics were published. The production, however, although so recent, has been so suggestive for future work that the time is undoubtedly propitious for taking stock of the tests at our command and for estimating their significance.

As in the history of educational measurement in general, one of the primary aims in the construction of tests of mathematical ability has been the establishment of standards. The admirable results already achieved by standardization in the elementary school subjects have naturally had considerable influence upon those, who desire to see measurement applied to high school work also. Consequently, we find that the first tests of mathematical ability were devised with a view to standardizing the simplest branch of the subject, namely ninth-grade algebra and that the technique employed has followed closely the methods adopted in the measurement of arithmetical abilities.

Thus in both cases use is made of the same two distinctive types of tests, based on methods of measuring that were first applied in psychology, but refined and improved in accordance with the development of the mathematics of the measurement of variables and also in accordance with the special purposes in view. One type gauges in terms of amount done primarily the ability to perform a uniform task. It emphasizes speed. Thus we have in first-year algebra, corresponding to the Courtis Standard Tests in arithmetic, tests of the same fundamental description by Monroe, Childs, Stromquist, and Rugg and Clark. Of these the two last-named writers have made by far the most elaborate inquiry into the achievement of pupils in ninth-grade algebra. They have had the assistance of several other workers in investigating in detail the many related questions that bear upon the problem; but it is the tests, which they have devised that have yielded the crucial evidence in determining what reconstruction is necessary. While they have carried out a much more extensive inquiry than these other investigators, their tests are in general the same in principles of construction, such differences as exist being due to the much more careful analysis of the mental functions involved in first-year algebra, the more comprehensive and representative abilities measured and the more economical devices applied.

The other type of test is the so-called difficulty, or development, or power test, which measures primarily how hard a task the pupil is able to perform. It consists of a number of problems, which present, as it were, a series of hurdles each exceeding the one before by a definitely determined increment of difficulty. Thus we have in algebra, corresponding to the Woody scales in arithmetic, Hotz's first-year algebra scales and Coleman's scale for testing ability in algebra. The makers of scales, while aiming at constructing a measuring rod having an absolute zero point, a definite unit of measurement and equivalent increments of difficulty, such as we have in our measuring instruments for physical phenomena like the yard stick or the thermometer, have in actual fact attained only approximations to such scales; yet while it is true that they are not perfect from the standpoint of the ideal measuring rod, these first-year algebra scales are nevertheless exceedingly valuable, if only because they furnish us with objective and relatively accurate measures. Thus the difficulty of each step has been calculated and fairly reliable standards of class performance have been determined.

The grading of the problems in the scales by Hotz and Coleman is based upon the proportion of pupils successfully solving each. Obviously where 80 per cent. of a group of pupils can successfully solve one problem and only 60 per cent. of the same group can successfully solve another, the latter must be a harder problem than the former. We can determine precisely the relative difficulty from our knowledge of the nature of the distribution of mental abilities in general in any random sample of human beings. This, however, is not the sole method that has been employed to grade problems in difficulty. In the first tests in algebra and geometry to appear, which were those published by Professor Thorndike in the MATHEMATICS TEACHER in 1914, the scaling of the problems was based not on the proportion of pupils solving each, but on the judgment of a group of two hundred teachers of mathematics as to their relative difficulty. In this instance Professor Thorndike was interested primarily in demonstrating to those who acted as judges the statistical procedure underlying the judgment scale in distinction from the performance scale. The former depends upon the principle described first in 1898 by Fullerton and Cattell, namely that equally often noted differences are equal.

To take a concrete case, where 50 per cent. of a group of experts judge that writing specimen A has more merit than writing specimen B and the remaining 50 per cent. hold the opposite opinion, then the two specimens are inferred to be of equal merit. We find that judgments are so distributed in cases where objective measurement is also possible, as in judging the relative length of lines. Where 75 per cent. of a group of experts state that writing specimen A has more merit than writing specimen B, and the remaining 25 per cent. support the contrary view, then writing specimen A is held to surpass writing specimen B by one unit of merit. In the scale so derived, the grading of the specimens depends upon the proportion of judges estimating one sample to be superior to another. In the instance under discussion a representative series of algebraic problems was collected. These were then compared each with every other by the two hundred teachers and so a scale representing the trend of their judgments constructed.

This method is properly applied in the case of English composition, drawing or writing, where the differences in merit *must* ultimately rest upon the judgment of trained jugdes; but the more objective method of measurement is preferable, wherever it is available. It is obvious that in algebra the performance scale should be used. We still can attribute some value to the judgment scale, however, even while granting its relative inferiority in the case of algebra, since it does yield an instrument for rough measurement at least.

The two main types of measuring rods—the standard test and the scale,—have their distinctive uses in accordance with their mode of construction. Thus the scale essentially consists of a series of problems, which although originally chosen with a view to securing as varied and inclusive a group as possible, finally retains only those which happen to be correctly solved by a proportionately increasing number of pupils. Thus the ultimate basis of selection is statistical. A test constructed in such a manner can scarcely avoid losing its original comprehensive and representative character. However varied the cases constituting the series in its first form may have been, in its final form it can hardly fail, as a result of this method of elimination, to lack many problems, with regard to which we need to know specifically, whether they can be solved by any children in a particular grade and precisely by what percentage or proportion. Measurement is necessarily inadequate, when in constructing instruments, the mental processes involved in learning are regarded as of secondary importance.

In this respect scales present a striking contrast to the best types of standard tests. In their case analysis is first made of the mental processes included in the field of study in question and examples of each type are incorporated. This implies a long and difficult task and it is a splendid achievement that we already have such an analysis of ninth-grade algebra. We had previously been unable to define in detail the ground covered as the formation of such and such specific mental bonds between such and such situations and such and such responses. At last now as a result of analytic study the mental associations and dissociations constituting elementary algebra have been determined. We have hitherto thought vaguely in terms of complex mental capacities such as ability to add, ability to multiply, ability to factor, or ability to solve an equation and the like, rather than of the more elementary bonds, of which those composite abilities are composed. The breaking up of those complexes into specific stimulus-response connections is of immense value; since successful teaching really consists in establishing these mental associations in pupils to the required degree of strength, in the best order and in minimal time. In other words, on this analysis all economical learning and teaching depend.

It is to the psychologists we owe the realization of the importance of analysis. Psychologists of every school—analytical psychology, functional psychology, behaviorism or what not—have this in common; they make an intensive analysis of the material they study—namely mental functioning. In ordinary life it is only too apparent that as a rule we fail to give any analytical consideration to the tasks with which we are confronted. These appear to us as units, literally en bloc. Such an attitude, however, leads to ineffective work and wasteful expenditure of energy. We can short-cut the lengthy and prodigal trial-and-error, hit-or-miss process of learning by a thorough-going analysis of the task, and of the most favorable attitudes for work. By such methods Gilbreth was enabled to treble the work of bricklayers and many other students of scientific management have discovered the sources of loss, waste and fatigue by similar means. Just as in industry to-day the efficiency engineer has by painstaking study reduced motor processes to their simplest elements, has determined the essential constituents of the skill in question, thereby eliminating useless and superfluous movements, has determined the best order of their formation and the amount of practice needed for perfecting each item in the total processes; so in the establishment of algebraic skill, analysis is an indispensable precondition and if the tests had accomplished nothing further, this would have been in itself a noteworthy contribution. It is a great step in advance to know the requirements of first-year algebra in detail from the standpoint of learning, from the standpoint of the pupil and the changes we are attempting to effect in him. We shall make advance in improving the teaching of algebra precisely in proportion as we succeed in making such analyses of its complex character.

The tests devised by Rugg and Clark stand out preeminently in this respect. They represent an analysis of ninth-grade algebra in to the unitary abilities, which compose it. They literally test one by one the strength of the bonds that together constitute ninth-grade algebra. A similar analysis of the work for higher grades still lies ahead of us.

Should economy of time demand a more restricted number of problems than we would have, if we included each distinct mental operation in our tests, the device of a spiral arrangement is used, by which several related operations are combined in the same test in cyclic order. This device first used by Judd in the Cleveland Survey arithmetic tests has been incorporated by Rugg and Clark in their algebra tests. To take a concrete example, the ability involved in removal of parentheses can be resolved into six fundamental types of mental associations. Accordingly, in the test to measure this complex capacity these six forms recur at regular intervals. It is worth noting some of the advantages of this device. One is that the score made by one pupil can be compared with that made by another. For example, if boy A attempts 24 problems and has 22 right answers and boy B attempts 12 problems and has 11 right answers, then we are justified in asserting that boy A has accomplished twice as much as boy B. While every mental bond or manipulation is not accounted for in the test so designed, at least the essentially new elements in learning to remove parentheses are included. Another advantage a spiral order offers is that it enables us to distinguish errors due to accident from errors due to ignorance or incapacity. Errors of the latter sort will recur at regular intervals and so be readily recognized. By this arrangement, too, we have a means of ascertaining not only errors common to the class as a whole; but also of locating individual errors as well and this entails the important consequence that thereby we are furnished with information indispensable for proper teaching emphasis both in the case of the class as a group and of the pupils individually.

It must be self-evident that this is one of the most vital uses to which tests can be put. When we have determined the class accomplishment and have compared it with those of other classes in other cities or with its own previous performance and so measured progress, interesting though that information may be, it has still to be interpreted. Should we find the results so obtained are unsatisfactory, this should merely instigate us to that further inquiry into the sources or causes of the condition revealed. The discovery of the latter is the greater discovery, which alone can suggest in what direction action is necessary. In addition to the help the tests can give in securing accurate measures of individual skill and of class performance and thus in improving current methods of classifying pupils, they can by diagnosing individual weaknesses and class weaknesses suggest radical alterations in methods of teaching.

Experimentation in arithmetic has demonstrated irrefutably that the recognition of individual errors and their special treatment is the direction along which we may look for the realization of the greatest economy of time. The diagnosis of individual weaknesses is the *first* step in the process of differentiating teaching to meet individual needs and as such forms the keystone in any time-saving program. The increase in efficiency resulting from a comprehensive scheme of adaptation of teaching to individual requirements cannot be fully foreseen, but it is practically certain that it will revolutionize teaching in algebra, as it has already revolutionized teaching in other fields.

What is still more worthy of note, standard tests can modify greatly methods of *learning,* for they are invaluable as furnishing motives for self-initiated activity, especially in connection with the more formal aspects of algebraic study. They act as incentives to zeal on the part of *pupils,* as well as of *teachers.* We all need such standards. We profit greatly by knowing our exact accomplishment, how we compare with our fellows and how our class compares with other classes, for it spurs us on and carries us over the routine work that must be mastered in elementary algebra. Teachers' marks expressed only teachers' opinions of their pupils' work. With the more objective, impersonal and reliable character of the scores derived from tests, pupils are conscious of a more powerful stimulus to endeavor.

They can be brought to realize that a certain standard of proficiency agreed upon by experts must be attained. The effect of hanging a Thorndike handwriting scale in the elementary classroom is the best illustration of the magical way in which standards work, leading the pupil in the case of a routine process to *drill himself,* until the required proficiency is achieved.

There is another purpose which the tests can serve. The statistical results already available are modifying markedly our ideas of what ought to be mastered in first-year algebra. Investigations of the mathematics in use in everyday life as also of the mathematical processes taught during the first year, which are essential for more advanced work in the subject, such as those undertaken by Rugg's collaborators, undoubtedly are essential in furnishing a clue to the new material that must be incorporated in the course of study. Nevertheless, we can reconstruct the content of our course wisely and scientifically, only if we know what has actually been accomplished by the material *now* in use. In other words we must first have accurate records of the results of present instruction.

Experimentation with new material and likewise experimentation with new methods of teaching depend for their success upon reliable means of gauging the effects of changed conditions. Standard tests and scales ring the death knell of mere guess-work and they hasten the demise of both rule-of-thumb methods and haphazard change. There will innevitably follow as a result of the more objective and accurate estimation of individual and class achievement, which they permit, the determination not only of minimal but also of maximal essentials, and thereby will be effected great increase of efficiency. For experimentation will reveal the optimum amount of drill, after which further practice would yield diminishing and unprofitable returns and so save time by setting limits of efficiency—maximal points, beyond which the pupil need not go and at the other extreme it will define minimal degrees of attainment, which we have a right to demand from efficient teachers of normal pupils. If this can be done in ninth-grade algebra and it has been in large part accomplished, it can also be done in more advanced work, in which we have every right to expect as fruitful results.

Standardization has also been begun in geometry, although nothing so thoroughgoing or ambitious as Rugg and Clark's algebra tests has been attempted. Four tests of geometrical abilities have been constructed by Minnick. In this investigation the capacities .fundamental to the demonstration of a theorem were selected for standardization, the subject-matter embraced being limited to the first two books of geometry. Three of these tests certainly measure abilities of first-rank importance; namely first the ability to draw accurate figures for theorems, secondly the ability to state the hypothesis and the conclusion of a theorem in terms of a given figure, that is to state exactly what is given and what is to be proved in the case of a specified proposition, thirdly the ability to select and organize facts to produce a proof. The fourth test which gauges the ability to recall related facts about figures, when one or more facts are given, appears to lack the definite purposiveness that characterizes true thinking and would seem to be less significant.

For these tests reliable standards have been calculated and they serve for geometry the same valuable ends that standard algebra tests have been shown to serve for algebra. It is by such intensive study of the abilities involved in geometrical work that improvement in the teaching of geometry will take place most rapidly. We are fortunate in being promised a scale for measuring pupils' ability to demonstrate geometrical theorems by the same author in a forthcoming issue of the *School Review.*

There are many other abilities essential to efficiency in elementary geometry besides those to which Minnick has directed his attention and some of these seem to me to be quite important. In particular, one group of capacities, which greatly influence success in elementary geometry is overlooked, namely the various functions included under intuitive grasp of spatial relations. The proof that two triangles are similar depends upon the power to apprehend immediately the spatial relations underlying the application of the principle of superposition. Individuals differ by native endowment as also from training in the dexterity with which they can accomplish this. A means of measuring the degree, to which it has been developed would be of great value in throwing light upon the causes of geometrical weakness or talent in any student. I have used the Thurstone Spatial Relations test, which in its original form measured the facility with which the pupil could put into practice the principle of symmetry, to estimate the ability to apply both principles, superposition as well as symmetry.

The need of an adequate measuring instrument for space intuition is still greater, when the student encounters for the first time solid geometry. We are familiar with the initial difficulties of handling tridimensional space. Here too great individual differences due to endowment as well as to practice exist. For diagnosis of special individual weakness a standardized test would obviously be of great service and it was for this reason I constructed a measure of the ability involved in intuitive grasp of tridimensional space. When we realize the important rôle spatial intuition plays in advanced mathematics, for example, in analytical geometry, we see that some method of gauging its strength is of significance not only for elementary geometry, but also for its prognostic power as regards future work.

In the *School Review* for October and November Irwin presents a *"Preliminary Attempt to devise a test of the Ability of High School Pupils in the Mental Manipulation of Space Relations."* Three hundred and thirty-seven pupils in three high school classes were examined in one test and three hundred and fifty-two in another. The material was not strictly geometrical; but the classification of errors is valuable for teachers of geometry.

It may be entirely fortuitous that standardization has not proceeded so far in geometry as in algebra; but no one can fail to recognize the comparative difficulty of isolating specific abilities in geometry, which is a necessary preliminary to designing suitable measuring instruments. Standardization is, indeed, a much harder task in geometry than standardization in algebra in virtue of the obstacles that confront us in resolving geometrical ability into the unitary abilities that compose it. It should be remembered that until recently we have been comparatively ignorant of the mental functions active in geometrical thinking. Whereas algebraic ability is readily differentiated into a number of easily distinguishable capacities, geometrical ability is more complex and presents less obvious lines of psychological demarcation and moreover the work done in arithmetic yielded helpful suggestions as to the realtive importance and the intrinsic significance of several algebraic functions. In elementary geometry this preliminary spadework had still to be done and much of it still lies ahead. Analysis was essential not merely to break up geometrical abiilty into its elementary constituents, but also to determine the efficiency and status of these unitary abilities in geometrical ability, to discover their interrelations and also their connection or lack of connection with other forms of mental capacity. It was essential to know whether there is an outstanding ability, which is the fundamental capacity in mathematical work and whether mathematical talent is such that it can function with approximately equal facility in relation to all kinds of material or whether it is in its very nature specialized and tied down to definite objects and situations. It was imperative for practical insight, that we should have experimental proof that a high degree of skill in algebra is accompanied by a high degree of skill in geometry or the reverse and that geometrical abilities correlate highly among themselves or otherwise. We grasp the full significance of any test, only when we do know its correlations with other tests. The process of determining these interrelations is a lengthy and laborious one; but ultimately it brings its reward. We cannot estimate the import of the reconstructions proposed in mathematics until such facts are discovered.

A fair amount of suggestive writing has been done along this line of analysis. Metzler and Stockard and Carleton Bell have made good suggestions for tests, although they have not attempted to satisfy the canons of rigorous scientific measurement. Indeed, Metzler's aim was merely to differentiate some of the more fundamental activities involved in geometrical work. Stockard and Carleton Bell, however, did design a test, for which as yet no standards have been published. It includes problems demanding the exercise of very varied mental functions for their solution. To it, therefore, there attach the typical drawbacks of the mixed scale. It cannot, for instance, provide us with a complete diagnosis of the gaps in the child's learning, the precise points at which mastery has not been reached. It has of course some significance, inasmuch as it can provide a rough instrument for classifying pupils in general geometrical ability; but it is not entirely satisfactory for the reasons stated. Anyone who looks forward to constructing geometrical tests would do well, however, to consider the fruitful suggestions that these writers have made.

A more systematic analysis of both geometrical and algebraic ability was published this year, based on strict methods of measurement. The primary aim of this investigation, which I conducted, was the measurement of mathematical intelligence and the obvious preconditions of success were first the analysis of the meaning of mathematical intelligence and second the construction of instruments by which its amount could be calculated.

Psychologists interested in the measurement of general intelligence had already blazed a trail through the jungle of mental measurements in the work they had accomplished. Binet and his followers,—Terman, Yerkes and others, had largely solved the practical problem of devising a measuring rod of general ability. Pearson, Spearman, Thorndike and their pupils had discovered the tools, by which an analysis of the nature of general ability into simpler mental functions could succeed. In measuring mathematical talent both problems arise, since neither mathematicians nor psychologists are agreed as to what the essential constituents of mathematical intelligence are.

After preliminary analysis, based on introspection of the mental processes involved in high school mathematics, tests were devised to measure as many of these as time would permit. These tests furnished a workable instrument for experimental investigation. For example, they served to determine the relative status of the various capacities measured in general mathematical ability. They were deliberately planned to gauge mathematical intelligence in children who had had five months of algebra and five months of intuitional geometry. My thought was that measurement of innate mathematical capacity was more urgent at this time than at any other stage in the pupil's high school career.

It would seem right in a democracy that no child of normal intelligence should be deprived of the opportunity of becoming acquainted with a realm of thought, which has meant so much for the advancement of science and of civilization as mathematics. Nor should he be deprived of a training in the use of the tools of quantitative thinking, except where he lacks the ability to profit by it. Both from a cultural and a practical standpoint knowledge of the fundamentals of mathematics should not be withheld, since some degree of mathematical mastery is necessary for advanced work in so many spheres. Nor should the child be considered a fit judge of what will be best for his future development, nor indeed should the parents in most cases. We need to safeguard the rights of the child, so that he is not steered away from the high types of work, of which he is capable.

Since the question of educational guidance is so critical, determining the future career of the pupil, allocating him to occupations of one level or another, we should have as reliable a gauge as the wit of man can devise. We require an independent estimate of mathematical capacity, which being as accurate and objective as we can make it, will serve to confirm or to call in question school marks, teachers' judgments and pupils' opinions of what they can perform. After two years' experience with algebra and geometry we would have the school marks in these subjects and the teachers' judgment of the pupil's ability. These are not necessarily the same, since the former are indicative of industry and other characteristics as well as of mathematical capacity. These together with an independent measure of mathematical ability derived from a group of tests would make it possible for a reliable decision to be arrived at whether further study would be profitable for the pupil. An independent estimate is particularly necessary in mathematics because routine processes play so prominent a rôle in elementary mathematical work. We hope that after reconstruction of ninth-grade algebra less time will be devoted to formal processes and more to thought-provoking tasks; but it is unlikely that ninth-grade algebra will not emphasize the mastery of the former to a considerable extent. Hence we need an instrument that will touch mental functions essential in more advanced mathematics and having little or no place in elementary work. This instrument should test other abilities than those tested by school examinations. It should supplement these and in this way provide us with a more adequate prognosis of mathematical skill and insight.

Consequently in the investigation undertaken the original abilities demanded by high school mathematics were discovered by pyschological analysis rather than the abilities that are largely the result of training. Of course, ultimately it is impossible to measure any ability apart from all training. We reach this limit only as an ideal. Where intelligence is primarily in question, however, problems requiring adaptation to new situations should predominate. From his measuring scale of intelligence Binet at one time eliminated all tasks presupposing instruction, but later he reintroduced them on the grounds that after all, if a child had had an opportunity of coming in contact with certain experiences, it is an index to his intelligence if they have left no trace or contrariwise.

Similarly in attempting to measure mathematical ability, while tests were for the most part of a novel character, demanding adaptation to numerical, algebraical and spatial material of a new sort; there were also included a few tests with which the pupil would certainly succeed better, if he had had previous training. Thus a test of algebraic computation was used as a rough gauge of the pupil's ability to manipulate abstract symbols and obviously the amount of practice received could greatly affect their scores.

In all seventeen tests were used. Of these five were tests of algebraic functions and a sixth, a reasoning test, measured an ability closely akin to algebraic ability. They were measures of ability in algebraic computation, ability to translate verbal statements of problems into algebraic symbolism, ability to match arithmetical progressions and corresponding formulæ, ability to analyze numerical data, to perceive a general rule implicit in them and to apply the principle so derived, and ability to make inferences in the case of purely symbolic material. There was also a test of skill in arithmetic problems. Five were tests of geometrical capacities, namely ability to make inferences with spatial data, ability to grasp intuitively spatial relations, bi-dimensional and tri-dimensonal and lastly ability to generalize in the field of spatial facts. Four were tests of language ability.

These were the well-known mixed relations or analogies test, the familiar logical opposites test, the Thorndike Reading Scales and the Trabue Language Scales. The inclusion of the language ability tests may at first consideration appear strange; but as our suspicion was and as the experimental results showed, the ability to conceive clearly the meaning of the words of any problem is as important an element in its solution as any connection it has with the abstract symbols of algebra.

Some of the results of the investigation are interesting in connection with current disputes. It has been advocated by several writers, notably by Breslich, that geometry and algebra should be fused as far as possible into a single subject. Our results, however, show that geometrical and algebraic abilities are no more closely related, no more dynamically connected with each other than with such superficially disparate abilities, as are measured by the analogies test, the logical opposites test, the Thorndike Reading Scales and the Trabue Language Scales. It is a striking fact that when the scores of the various geometrical, algebraic, and language tests are combined respectively into three groups, the coefficients of correlation between these composites are practically identical. The crude coefficient for algebra and geometry was .38, for algebra and language ability .42 and for geometry and language ability .41. The corrected coefficients, which probably give the most true amount of interdependence between the mental functions in question yielded similar results. The coefficient for algebra and geometry was .49, for algebra and language ability .56 and for geometry and language ability .56. Apparently algebra and geometry are in no special degree closely akin.

The results further emphatically indicate that a command of the vernacular is an important precondition of success in mathematical work. The ability to comprehend words, to grasp accurately the significance of a problem is as vital an element in its solution as the manipulation of algebraic symbols. Too little heed is paid to this aspect of mathematical learning. Much more time should be given to clarifying quantitative thinking as expressed in language. These results fully substantiate Rugg and Clark's contention that much more effort should be devoted to translation, not only from English into algebraic symbolism, but also from algebraic symbolism into English. Only by such exercise will the abstract entities we deal with in algebra become familiar tools in the control of the learner.

Another result worthy of note, although merely confirming common sense, is that no single test can be a sufficient index to mathematical intelligence. Psychologists have usually acted on a different assumption and it is an interesting finding that not one of the seventeen tests used correlated at all closely with the best available measure of mathematical capacity. If there is any one dominant mental function in mathematical ability, it is that which is common to those tests, which probe mathematical intelligence most effectively and this seems to be the ability to react to parts of a situation, in other words to analyze and think selectively. Tests having this characteristic quality in marked degree form the best index to the presence of mathematical talent, although it certainly involves many other functions indispensable for success with the subject.

This investigation also showed that great individual differences in mathematical skill and insight exist. The presence of a very wide range of ability was never more clearly demonstrated and the need for classification in accordance with capacity was never more apparent..

A still more striking result and one that made a great impression upon my mind was the evidence it furnished of the ability of the pupils to attack much harder tasks that we commonly assign to them. We seem prejudiced against the child's ability. The extreme instances of sheer incapacity influence us unduly. Again the need for classification according to inborn power shows itself to be urgent. When the question of reconstructing the content of the elementary courses in mathematics is before us, it will be necessary for many of us to take a less depreciatory view of the capacity of the pupil and to ask ourselves whether we have a correct conception of what he can do, rather than to doubt the advisability of introducing more advanced material earlier to the student. I incline to agree with Thorndike that there is no kind of reasoning a child of eleven cannot do, provided the material lies within the realm of his experience. Indeed, it would be easy to find a child of eleven, who could do more efficient mathematical reasoning than many adults. Opportunities for the gifted should not be lacking, nor should we take so low a view of the ability of the average child.

By determining the functional interdependence of the traits measured by those tests and their relative importance in mathematical intelligence, it was possible to select a team of six tests, which while possessing maximum diagnostic power was also easy of application and economical of time. The prognoses derived from combining the results of these tests closely correspond with the actual future achievements of the individuals tested. In the case of a group of 61 pupils in Horace Mann School, an hour and a half of time spent on the tests gave a coefficient of correlation of .82 with future mathematical accomplishment.

Thus by the aid of this sextet of tests we can predict not with perfect precision, but with a known degree of accuracy the capacity of the pupil to undertake more advanced work. By its use we are able to classify pupils on a more reliable basis, so improving current conditions, by eliminating from the mathematics classes those unfit for further training and selecting those of superior ability for special treatment.

Much more scientific work remains to be done along the line of this investigation. We need tests to measure innate algebraic intelligence, before the child has had any algebraic training whatever. We should have a test or a group of tests of such a character applied the first day in the ninth-grade algebra course, so that we could from the outset group pupils more appropriately, in consonance with their powers and their promise as students of the subject. The same holds good of geometry. For it, too, we need a team of tests or a single test of high diagnostic power, which presuppose no previous knowledge with the subject, and which could provide us with an adequate measure of innate ability in that field.

Prognostic tests to improve classification in our larger high schools, standardized tests for more advanced algebra and geometry—these are the tasks that await our attention. The work will be difficult; but the results will be proportionately precious. Surely in no other subject are rule-of-thumb methods more ridiculous than in the teaching of mathematics. Surely in no circumstances have we a greater right to expect helpful assistance in measurement than from a group of people trained in quantitative thinking, with much of the equipment that is necessary to carry on the work. Mathematics teachers should lead in this movement to replace subjective and unreliable school marks by objective, unbiased and accurate methods of measuring. We would thereby not only be aiding in the improvement of the teaching of our own subject, but we would also be making of education a science and playing a part in one of the great movements of modern times.

We can scarcely grasp the full significance of the introduction of scientific measurement in education. We can but dimly discern its probable results. When clairvoyance and divination were for the first time in history replaced by science, by the inductive experimental method, how few were conscious of the momentousness of the event that had taken place! How few realized that a turning point in the history of mankind had been reached. In like fashion we apprehend but faintly that a new era has begun in the records of the human race. The birth of physical science was a world-shaping event, inasmuch as it gave control over the material universe. The application of measurement and experimentation to mental products is as remarkable an occurrence, for it literally means control of the human machine. The last factor in efficiency to be controlled, the human factor will in all likelihood yield the richest results of all. It seems only yesterday that we ignored the fact that a few unnecessary movements in an activity like sewing adds much to the bodily energy consumed in a day, so we now are for the most part unaware of the wastefulness of our methods of habit-formation in mathematical work. As soon as effort is devoted to changing this state of affairs, we shall be magnificently rewarded in the saving in human strength that will result. At this time in the world's history, when the minds of men were never more malleable, when far-reaching adjustments in our social and in our educational institutions are in progress, we should be eager to play a part in a revolution still more fundamental, more invasive and pervasive than any other—the downthrow of opinion and the rise of science in the schoolroom. When we recall how long it took before we had precise instruments of measurement in physical science and how much is expended yearly to maintain bureaus of weights and measures, we shall not grudge the cost of obtaining educational measuring rods, although the magnitude of the task confronting us may temporarily give us pause. It is well to remember, however, that in no subject are circumstances more favorable, in none should progress be so rapid as in mathematics. The work already accomplished although fragmentary, has already proved its value and gives promise of richer results. Undoubtedly the tests at our disposal have much significance for the mathematics teacher. They have begun the detailed analysis of mathematical abilities, they have rendered possible improved methods of classification, and diagnosis and prognosis of individual ability can now be made by their help. Analysis, classification and diagnosis—these three outstanding ends they serve; and the greatest of these is diagnosis.

BIBLIOGRAPHY.

CHILDS, H. G. "The Measurement of Achievement in Algebra." Bulletin, Extension Division, Indiana University, Vol. 2, No. 6. pp. 171–183.

COLEMAN, W. H. "Scale for Testing Ability in Algebra." Superintendent of Schools, Bertrand, Nebraska.

FULLERTON, G. S., and CATTELL, J. McK. "On the Perception of Small Differences." Publications of the University of Pennsylvania Philosophical Series, No. 2, The University of Pennsylvania Press, Philadelphia, 1892.

HOTZ, H. G. "First Year Algebra Scales." Teachers College, Columbia University Contributions to Education, No. 90, New York City, 1918.

IRWIN, H. N. "A Preliminary Attempt to Devise a Test of the Ability of High School Pupils in the Mental Manipulation of Space Relations." *School Review,* Vol. 24, October, November and December.

METZLER, W. H. "Problems in the Experimental Pedagogy of Geometry." *Journal of Educational Psychology,* Vol. 3, p. 545.

MINNICK, J. H. "An Investigation of Certain Abilities Fundamental to the Study of Geometry." The University of Pennsylvania Press, Philadelphia, 1918.

MONROE, W. S. "A Test of the Attainment of First-Year High-School Students in Algebra." *School Review,* Vol. 23, pp. 159–171.

ROGERS, AGNES L. "Experimental Tests of Mathematical Ability and Their Prognostic Value." Teachers College, Columbia University, Contributions to Education, No. 89, New York, 1918.

RUGG, H. O., and CLARK, J. E. "Standardized Tests and the Improvement of Teaching in First-Year Algebra." *School Review,* Vol. 25, pp. 113–132 and 198–213.

"A Cooperative Investigation in the Testing and Experimental Teaching of First-Year Algebra." *School Review,* Vol. 25, pp. 346–349.

"The Improvement of the Ability in the Use of the Formal Operations of Algebra by Means of Formal Practice Exercises." *School Review,* Vol. 25, pp. 546–554.

"Scientific Method in the Reconstruction of Ninth-Grade Mathematics." The University of Chicago Press, Chicago, 1918. Supplementary Educational Monographs, the *School Review,* Vol. 11, No. 1.

STOCKARD, L. V., and BELL, J. Carleton. "A Preliminary Study of the Measurement of Abilities in Geometry." *Journal of Educational Psychology,* Vol. 7, p. 567.

STROMQUIST, C. E. University of Wyoming, Laramie, Wyoming.

THORNDIKE, E. L. "An Experiment in Grading Problems in Algebra." MATHEMATICS TEACHER, Vol. 6, p. 123.

GOUCHER COLLEGE,
BALTIMORE, MD.